Fiction and Psychological Insight

Kathleen Stock

It is not unusual for a reader of a novel, especially that of the nineteenth century variety, to assume that, in reading, she is acquiring important insights into human beings. Yet philosophers have often found this assumption problematic. Most agree that fiction can be a source of psychological understanding, either explicitly, via psychological descriptions of characters, or implicitly, via the construction of psychological character portraits. However, there is disagreement about the importance of fiction's potential contribution in this area.

Some have suggested that the psychological information presented in a work of fiction could not reasonably strike a reader as true without the reader having come across it already in some other non-fictional context. Jerome Stolnitz represents this view when he writes: "Art, uniquely, never confirms its truths. If [on reading Jane Austen] we find that stubborn pride and ignorant prejudice sometimes keep attractive men and women apart, we find the evidence for this truth about the great world in the great world" (1992: 198; see also Diffey 1995: 210). Meanwhile, Noël Carroll suggests that, in the moral realm, which presumably includes the realm of moral psychology, "if… learning is a matter of the acquisition of interesting propositions heretofore unknown… then… there is no learning when it comes to the vast majority of narrative artworks" (1998: 141). In contrast, others have cast fiction as a potentially 'self-sufficient' source of psychological understanding (Robinson 1995; Conolly and Haydar 2001: 119).

As already intimated in Carroll's remark, a second prong of the attack has it that the psychological information provided by fiction cannot be genuinely *interesting*; that, at best, it is made up of 'truisms' (see also Stolnitz 1992: 194). Others deny this, emphasizing the importance of the information acquired (Graham 1995; Robinson 1995; Conolly and Haydar 2001).

It seems obvious that fictions can be a 'self-sufficient' source of interesting psychological insights, and it seems curious that some writers have thought otherwise. One contributory factor here might be a certain picture of what it is to acquire psychological understanding, according to which such understanding is inductively acquired through extrapolation from evidence. On this view, where p is a proposition expressing some piece of psychological information, one comes to recognize p as true by encountering, either directly or via testimony, partial evidence for p, with the degree of confirmation for p, increasing, at least initially, in proportion to the number of occasions upon which such evidence is encountered.

Such a picture might appear to validate both the thought that a work of fiction cannot be 'self-sufficient' with respect to any psychological information disseminated, and that a fiction can provide knowledge only of uninteresting truisms. One who supposed that psychological information was acquired inductively, via direct or indirect exposure to relevant evidence, might be tempted to treat a psychological portrait or description delineated in a fiction as a kind of testimony – that is, as a form of second-hand

evidence of the phenomenon picked out. If this were so, then the mere fact of its single appearance in a fiction could not be enough, on its own, to show the reader that it picked out a genuine phenomenon 'in real life;' the reader would have to be exposed to additional appearances before she could assume this with any reliability. (Analogously, neither would a single report of cold fusion legitimize the assumption that cold fusion was possible.) In fact, assuming that one treats a psychological description expressed in a work of fiction as testimony *about* the phenomenon in question, rather than as some kind of first-hand evidence *of* that phenomenon, additional complications about the reliability of the information presented seem to be raised, along the lines of those raised for testimony generally.

At this point, one might take one of two positions (or, like Stolnitz, take both). Perhaps motivated by worries about the reliability of the testimony of authors concerning psychological matters, one might deny that a psychological portrait or description in a fictional work could count as *any* kind of evidence for a psychological truth – authors, after all, are not generally trained as psychologists and are all too prone to imply contradictory statements (Stolnitz 1992: 196). Or, less counterintuitively, one might acknowledge that fiction can be, and often is, a source of psychological understanding for a reader (Stolnitz 1992: 193), in which case, at least two things would follow. First, any information presented in a work of fiction must be recognized *as information*, and must be recognized as true based on evidence *already* encountered by the reader prior to reading that work; this entails that works of fiction cannot present, on their own, wholly new truths. Second, any information thus presented would tend not to be of great complexity or interest, since the more complex and specific a psychological portrait or description expressed in a fiction, the less likely one has already found evidence for it elsewhere.

This view of psychological understanding as a kind of inductively acquired understanding appears to implicitly ground several authors' skepticism about fiction's potential for meaningful instruction. Stolnitz claims: "None of its truths are peculiar to art. All are proper to some extra-artistic sphere of the great world" (1992:198). Terry Diffey writes "we cannot learn from a work… unless we already know that… the world is as the work shows it to be" (1995: 210). Meanwhile, Carroll's apparent presupposition that psychological understanding must be empirically acquired has been seized upon and criticized by certain other commentators (Conolly and Haydar 2001: 121).

I shall not attack the premises of such writers directly. Rather, I shall reject their conclusions, arguing that there are certain psychological depictions which, when presented in a fiction, in virtue of their very intelligibility to the reader, reveal their nature as possibilities of human experience, thereby bringing the reader to new psychological understanding of a propositional kind. In these cases, the reader does not need any prior (or subsequent) evidence of such phenomena in order to see that they are genuine possibilities for human life – states of mind actual human beings might have – and so the psychological understanding acquired by the reader through exposure to such portraits is not acquired inductively. This means that there is restriction neither on the novelty nor on the interest of the information potentially acquired in such a manner.

First, though, I describe some of the ways in which a fiction may make certain actions of its protagonist intelligible. In later sections, I show that these cases potentially provide psychological understanding, not just of fictional characters, but of real people as well.

Fictions and the Intelligibility of Actions

Uncontroversially, a fiction can make an action of a character intelligible. On what one might call a 'weak' or 'formal' conception of intelligibility, a fiction makes a given fictional action (or mental state) intelligible simply by filling out aspects of the psychology of the fictional agent, showing how the action in question coherently relates, in terms of means and ends, to other aspects of her mental set. For instance, a fiction might make George's joining the circus weakly intelligible, by making it clear that he wants to become a lion-tamer, and that he believes joining the circus will enable him to do this. Perhaps it also reveals that he wants to become a lion-tamer in order to outdo a rival, that he believes becoming a lion-tamer will enable him to do this, and that he wants to outshine his rival in order to impress his beloved… and so on. A minimal requirement upon making a fictional action intelligible in this weak sense is to start by describing some goal of the fictional agent, and then to attribute to her a belief that the action in question will achieve, or contribute to achieving, the goal. The resulting description of a belief and an attitude will constitute the agent's reason for so acting (Davidson 1980a: 3-4). One might then provide a context for this reason, by showing how it fits with further goals she has and with her beliefs about how the action will achieve such goals.

There is an obvious sense in which fictions can do this. For instance, Nabokov's *Lolita* makes the kidnapping and drugging of a child formally intelligible by showing how those actions cohere with the protagonist's desire to seduce the child, and with his belief that kidnapping and drugging her is a means of doing so. Similarly, Toni Morrison's *Beloved* makes infanticide weakly intelligible by furnishing an account of the surrounding states of mind which might accompany such an act.

However, simply showing how a given fictional action or mental state coheres with other mental states need not on its own say anything very informative. Finding intelligible, in this weak sense, a fictional character's action A, in terms of some prior desire D_1 to achieve a certain goal G_1 – say, the act of procuring a saucer of mud because one desires it (the example is from Anscombe 1963) – might just be a matter of attributing to the character another desire D_2 for some further goal G_2 (say, the desire to rub mud into one's face), and a belief about how satisfaction of D_1 is a means of satisfying D_2. For each new desire so posited, such (weak) intelligibility could be a matter of explaining *that desire* in terms of yet another desire, in much the same way. The positing of such further desires, along with relevant beliefs about the relationship between the relevant goals, could go on for some time without A being made intelligible to the reader in an interesting sense. To anticipate: it would be of no use merely to be told, for instance, that the character procures a saucer of mud with which to rub mud into her face, if one did not also understand *why* she might want to rub mud into her face. Additionally, to be told, say, that she wants

to rub mud into her face because she wants her face to be the same color as her wardrobe, would not help matters. Something else is needed, and it is not just the positing of some further desires and beliefs *per se*.

A distinction therefore should be drawn between weak or formal intelligibility and substantive intelligibility. A central way to make a given fictional action *substantively* intelligible is not simply to cite certain mental states which cohere with it, but also to show that the mental states with which it coheres are *themselves* substantively intelligible. I shall clarify this by focusing on two ways in which an author might show the mental states of a fictional character (and thereby any actions of hers motivated by such states) to be substantively intelligible.

First, take the case of desire. The citing of a desire is a central way for an author to rationalize a certain act of a fictional agent: she did A because she wanted G (and she believed that A was a means of obtaining G). But in order for this to work as an interesting rationalization, the desire itself has to be made substantively intelligible.

A central way in which a reader comes to find a fictional character's desire to achieve a certain goal substantively intelligible is via the fiction's making manifest its relation, ultimately, to a 'desirability characterization' (Anscombe 1963: 70-2); that is (or so I shall claim), to a goal which the reader judges as desirable, ceteris paribus. One understands, in this stronger sense, a character's desire D for goal G_1, if one sees that the character wants to achieve G_1 in order, ultimately, to achieve G_n, where G_n is a goal one judges to be desirable, ceteris paribus.

Take again, for simplicity's sake, Anscombe's unusual case of the desire for a saucer of mud, itself made weakly intelligible by its connection to the agent's desire to rub mud into her face. Both of the desires cited here might become substantively intelligible to a reader R upon R's realizing that the character who desires this, does so in order *to improve the condition of her skin*. Such a realization would make the action substantively intelligible to R where R judged that, other things being equal, improving the condition of one's skin is a desirable thing to do. In contrast, to revert to the former example, the positing of the further goal *to make one's face the same color as one's wardrobe* would not make intelligible to R a character's wanting a saucer of mud in the same way, since normally R would have to struggle to see how such a desire could be connected, immediately or mediately, to any desirable end.

Fictions, then, can make an apparently unintelligible desire of a character comprehensible, not only by furnishing a coherent account of surrounding mental states and behavior, but, more interestingly, by showing that the desire falls under, or is grounded in a further desire which falls under, a desirability characterization. In *Lolita*, for instance, Humbert wants to seduce the child Lolita, and believes that kidnapping and drugging her is a means to doing so. This desire, perhaps only formally intelligible to many when so baldly stated, is made substantively intelligible for the reader when the desire is related, either directly or indirectly, through Nabokov's insidious art, to some desirability characterization: perhaps, a desire to regain feelings of one's sexual awakening; or a desire for sexual possession of a person with at least some of Lolita's physical characteristics (her "glowing" skin, "the silky shimmer of

her temples shading into bright, brown, hair" (Nabokov 1995: 41) and other physical characteristics associated with beauty and health are constantly emphasized). Making this desire intelligible renders Humbert's subsequent behavior, motivated by this desire, also intelligible. (The fact that *Lolita* as a whole alternately invites and evades such interpretations is consistent with my eventual claim that the work makes intelligible certain reasons for acting as Humbert does, whether or not it is true in the fiction that these are *Humbert's* reasons for so acting). **[Comment: Should the word 'insidious' be used to describe Nabokov's art? The word means: 'proceeding in a gradual, subtle way, but with harmful effects.' Do you wish to imply that Nabokov's work was harmful?]**

Of course, for many mundane cases, there is no need for a fiction to explicitly reveal the desirability characterization of a given action, since the reader will readily anticipate that characterization. I am concerned, rather, with the sort of fictional action which would, on a condensed characterization, be unintelligible to most readers, without the fiction's (often prolonged and perhaps implicit) articulation of the desirability characterization under which the action falls.

Some clarifications: where a desire is made substantively intelligible to reader R by its ultimate relation to a goal which R judges to be desirable, R need not be concerned to pursue this goal herself. This is part of the point of the ceteris paribus clause; it indicates the possibility that R might have other goals which she deems more valuable than the goal cited in explanation. In this case, the fact that she herself does not pursue that goal as an end does not show that she does not value it.

Secondly, that a desire D of a character becomes substantively intelligible to R by its being related, ultimately, to a goal which the character judges to be desirable, is consistent with R's finding the goals to which D is more immediately related positively undesirable, even reprehensible. This is shown by the *Lolita* case, but the point can be made more clearly in relation to a simpler example. Say that a fictional character on a train beats up a fellow passenger in order *to spread blood all over the carriage* (I owe this example to Peter Goldie). Perhaps we can find this substantively intelligible if it is related to the further goal of *making a pleasing pattern*, an end which we understand, if not as enormously important, then at least as having some value, ceteris paribus. In this sense, we can come to understand the action. But this is not to say, of course, that we think it a good thing to spread blood all over the carriage. For, quite properly, we also judge that any value the action would accrue, when understood as the act of making a pleasing pattern, is outweighed or even negated by other consequences of the action when it is fully described. (In another sort of case, one might acknowledge the desirability of a goal towards which an action is ultimately aimed, thereby acknowledging the action as substantively intelligible, while judging that the action is nonetheless a bad thing, because one judges that it is not in fact likely to achieve the goal in question.)

Acknowledging the former sort of case means that there may yet be a sense in which a substantively intelligible desire of a character, in the sense I have described, is nonetheless unintelligible: namely, insofar as the reader cannot understand the character's prioritization of the desire over other

considerations, or cannot understand how the value for the character of achieving a given desire is not outweighed or negated by other consequences. Though there is not the space to explore the matter here, it seems that a fiction may also make a desire, or action motivated by such a desire, intelligible in this richer sense. In this paper, however, I shall largely focus on substantive intelligibility in the more limited sense just delineated (though I shall briefly return to the richer sense).

Finally, I claim only that being related to a desirability characterization is a *sufficient* condition for the substantive intelligibility of a desire, not that it is a necessary one. This means my position is untroubled by cases where fictional actions, although apparently substantively intelligible from a third-person perspective, are yet genuinely pursued *sub specie mali* – such as, perhaps, the activities of Satan in *Paradise Lost,* whose dictum is "Evil, be though my good" (see Anscombe 1963: 75; Velleman 1992:18; Dancy 1993: 6). For, if we must accept such cases at face value, it can be allowed that they are substantively intelligible in a way different from the cases with which I am concerned. Of course, the rejoinder may be that the substantive intelligibility of an action generally has *nothing* to do with any motivating desire being connected to a desirability characterization. I fail to see, however, how such a claim could be maintained, both when confronted by the intuitive force of the sorts of literary cases already discussed (that is, cases where an otherwise incomprehensible fictional action is made intelligible by its perceived connection to recognizable values), and, perhaps more centrally, when confronted by the characteristic way in which ordinary people try to make sense of others (real or fictional) by positing possible motivations for their acts in terms of goals readily understood as desirable. (For instance: 'why do you think she hit him?' 'well, perhaps she wanted to let off steam.')

Thus far, I may have given the impression that, at least for some cases, *all* there is to making a given action of a fictional character substantively intelligible is to show that the action is related, mediately or immediately, to a desire for an end which falls under a desirability characterization. This is misleading, insofar as, in such cases, it also must be possible to see how that particular action could be counted as a means of achieving the end in question. In other words, the belief of the character that the action in question will achieve, or contribute to achieving, the relevant goal picked out by her desire, must be shown to be intelligible too.

Normally, the fit between a character's desired end and any action supposed to be motivated by it need not be made explicit, since it is obvious how one relates to the other. Nonetheless, that there *must* be such a fit is not a given; it is conceivable that in some work of fiction (probably, a bad one), a character might be described as doing one thing in order to achieve another thing, where it is unclear to the reader how the former could possibly be counted by the character as a means of achieving the latter (Anscombe 1963: 35-6). For a reader to find intelligible a fictional agent's belief B that an action A is a means of achieving some goal G, she must be brought to understand what features A has, such that they might be conceived of by the agent as contributing to or consisting in achievement of G. It does *not* require the reader to believe that an action such as A *is* a means of achieving G; there has to be room for the reader to

judge the fictional agent's reasoning as faulty without the agent's belief thereby collapsing into unintelligibility. However, if the reader does judge B to be false, so that she judges A is not actually a means of achieving G, then in order for B to count as intelligible, the reader has to be able to see how the agent might have (mistakenly) arrived at B, given the agent's epistemic situation. For instance, in *Madame Bovary*, when Charles Bovary conceives of the operation on lame Hippolyte as a means of enhancing his prestige and so also Emma's love for him, the prescient reader knows that this is impossible; notwithstanding, she is not required to believe that the operation will be such a means, but only to be able to see how Charles could think that it might be.

This latter observation allows us to shed some further light on what I earlier called 'weak intelligibility.' For it turns out that making an action weakly intelligible is identical to positing a motive for an agent's action in terms of a desire or other pro-attitude towards some eventual end, where one can see how the agent might reasonably believe the action is a means to that end, but cannot understand what might be valuable about the end in question. In other words, the fit between the action and the posited desire is intelligible, but the desire itself does not fall under any identifiable desirability characterization.

So far I have been concerned with the intelligibility of desires. I now wish to turn, albeit briefly, to another sort of mental act with obvious relevance to the rationalization of action: moral judgment. As it is with a desire, a central way of making a fictional character's moral judgment (and thereby any actions motivated by that judgment) substantively intelligible is to show that such a judgment falls under a desirability characterization which gives it a point in relation to recognizable human interests. (In such cases, it is also important, as before, to make it apparent how the action in question could count as a means of achieving the desirable end picked out by the moral judgment, a point I shall take for granted in what follows).

For instance, Toni Morrison's *Beloved* makes intelligible a mother's murder of her child, who is otherwise facing slavery, by allowing the reader to conceive of the act as issuing from an intelligible moral imperative, falling under a desirability characterization – namely, "to protect every bit of life she had made, all the parts of her that were precious and fine and beautiful" (Morrison 1987: 163; see also Carroll 1998: 155). Under this description the moral judgement implicit in the heroine's action has a manifest relationship to what is recognizably worthwhile in human life, and is substantively intelligible as such. Other moral judgments issuing in action may need their point unveiled more explicitly: in Evelyn Waugh's *Brideshead Revisited*, for instance, Julia Flyte gives up her beloved Charles Ryder due to a growing awareness of the demands of her Catholic faith, and the recognition that she cannot live in sin with him, since she cannot bring herself to "set up a rival good to God's" (Waugh 1962: 324). The reader can only grasp the point of the moral imperative behind this act of self-sacrifice once she has grasped how Julia's sense of her own integrity centers on a truce with God that represents both reparation for past impieties and solidarity with her devout family, especially in light of the emotional impact of her formerly

recalcitrant father's deathbed conversion. That this is so is made clear (thought not explicit) in the record of Julia's final conversation with Charles.

The claim I endorse here reflects something of Philippa Foot's point that the action of "clasping the hands three times in an hour" could not be intelligibly thought of as a morally good action unless a 'special background' were built into the thought experiment which revealed its point relative to some aspect of human flourishing (1967: 92). In other words, a moral concept does not attach meaningfully to an action or situation simply in virtue of its employer's feeling some capricious sentiment of approval or disapproval towards that action or situation, which on another occasion she equally well might have withdrawn. Rather, its application has to be recognizably keyed to the promotion (in the case of 'positive' moral concepts) or the undermining (in the case of 'negative' moral concepts) of recognizable human interests. (Foot's insistence that moral concepts are keyed to shared human interests is endorsed by those of otherwise very different views on the nature of moral claims: cognitivists like McDowell 1994: 83-4 and noncognitivists like Blackburn 1984: 197.) This being the case, one instructive role for literature to play is that of showing how an apparently idiosyncratic application of a moral concept in a character's judgment of a situation can nevertheless be, in that particular fictional context, substantively intelligible to the reader.

Recall that to find a moral judgment intelligible in virtue of its connection to some ultimate goal one judges desirable is not necessarily to endorse the judgement itself. One can see why *Beloved*'s Sethe thinks it is right to kill her child, because one sees that her intended end is valuable, without necessarily agreeing that it *is* right that she killed her child. For instance, one may demur on the grounds that her action is not in fact a means of achieving that end, or on the grounds that other ends, whose achievement her action precludes, are in fact more valuable. For this reason, I reject the suggestion that, where a fiction makes a given moral response (substantively) intelligible, it thereby involves the reader temporarily 'entertaining' in imagination that attitude or response. (For example, Matthew Kieran interprets the reader's finding the protagonists of *Brideshead Revisited* intelligible as a matter of "entertaining the moral perspective of Catholicism" – Kieran 2001: 32.) To find substantively intelligible a moral judgement requires only that one be brought to 'see the point' of the judgement; to see it as related ultimately to a goal one judges, other things being equal, as desirable (as my discussion of the *Brideshead* case shows). This is not the same as temporarily entertaining or endorsing the judgement oneself. To find Julia Flyte substantively intelligible, the reader does not have to think or imagine, even temporarily (assuming that this is possible), that Julia is *right* to leave Charles, or that living with Charles *would actually* constitute "setting up a rival good to God's." One only has to see how *she* might come to think that, and to see what, from her perspective, might appear valuable about it as such. This, as my argument suggests, is a matter of recognizing the structure of her mental states and behavior, both as constituting a largely coherent pattern, and as having a recognizable point, relative to her epistemic position, even if one simultaneously rejects elements of that epistemic position as false or wrong.

One might object that to find Julia Flyte *wholly* intelligible (not just in the sense that one understands her action in the light of some goal judged desirable, ceteris paribus, but also in the richer sense, acknowledged earlier, that one understands how she could prioritize that goal over others also judged desirable) one must indeed be able to imaginatively adopt her evaluative perspective, if only temporarily. It seems to me, however, that in order to find a character's behavior intelligible in this richer sense, one is required, not to *imaginatively* adopt her evaluative perspective, but rather to *actually* adopt her evaluative perspective, that is, to be brought to see that whatever goal is in question *is* more desirable than others the character has. This, unlike imagining, is not a temporary state. In fact, it is not clear to me what 'imaginatively adopting an evaluative perspective' amounts to, if it does not involve actually adopting the evaluative perspective of a fictional character. However, I do not have the space to pursue this matter here.

Here then are two ways in which aspects of a character's mental states, and hence any actions motivated by those states, may be made substantively intelligible to a reader. Of course, these are not the only ways a fictional action may be made intelligible. An action might be made substantively intelligible in terms of some prior emotional response, for instance. (For a detailed elaboration of what might be involved, see Goldie 2000.)

The rationalizations under discussion need not be detailed explicitly. Usually, and certainly in the most interesting cases, the reader must infer the mediate source of a character's motivation, and its perceived value, under the guidance of what *is* made explicit in the fiction. An implicit rationalization may be countered by pronouncements of the fictional agent herself; just as in life, it is possible for an agent to be self-deceived about her motivation while her behavior betrays it to others. After all, in order for a reader to be brought to understand a fictional agent's motives, it is not required that the agent herself be depicted as aware of those motives.

It is also important to reiterate that a reader may understand a character's reasons for doing A in the sense in which I am interested without necessarily judging that the character has good reason to do A. I have deliberately focussed on examples where this is the case: for instance, one need not judge that there are good reasons for Humbert to kidnap Lolita or for Sethe to murder her children in order to judge such actions substantively (and not merely formally) intelligible. One makes such substantive judgments by finding the motivations of these fictional agents substantively intelligible. In such cases, either one rejects the attitudes of the agent which prompt the action in question, as corrupt, ill-disciplined, disproportionate, or otherwise inappropriate, despite their traceable relation to some goal the reader judges as good, other things being equal, or one rejects as false the agent's beliefs that her actions are a reliable means to achieve her ultimate ends. Indeed, often the author deliberately encourages such rejection. For instance, in *Lolita*, Humbert's attempts at justifying his treatment of Lolita are undercut by means of casual (from Humbert's perspective) yet revealing (from the reader's perspective) remarks he makes about her: her manifest boredom, her bouts of weeping, and so on. Of course, a work of fiction can lead the reader to see

the motivating reasons of a character as good ones. Yet it is one of the great functions of fiction, and especially of the novel, to reveal how a character's perspective on the world may come apart from what (fictionally) is the case; that is, to demonstrate that what the fictional agent ultimately holds valuable, though perhaps valuable ceteris paribus, should not be pursued in the particular context in which it is being considered.

Fictional Reasons and "Real Life" Understanding

One might wonder what any of this has to do with understanding of real life cases. After all, one might protest, fictions are about fictional entities, not real ones. How can a work about fictional characters and events have implications for how we might interpret real people?

Behind this worry lies the picture alluded to above in my introduction: that psychological understanding is acquired inductively. However, this is the wrong way to look at things, at least with respect to the substantive intelligibility of the sort of cases under discussion.

I have argued that a fiction may make a character's action substantively intelligible by (1) establishing a connection between the action in question and a desire/ moral judgment of the character's which falls, immediately or mediately, under a desirability characterization; and (2) further establishing the reasonableness of the character's belief that the action in question counts as contributing to the goal invoked in the desirability characterization cited in (1).

The successful execution of (1) forces the reader to acknowledge that the action n question is motivated in the light of some worthwhile or desirable goal, ceteris paribus. Now, one could not acknowledge that some goal G was a worthwhile one for humans to pursue, other things being equal, yet at the same time maintain that no human being ever could be motivated to pursue G. In other words, in acknowledging that a goal is a worthwhile one for humans to pursue, ceteris paribus, one *just is* acknowledging that some actual human being might be motivated to pursue it.

Meanwhile, the successful execution of (2) tells us that the character could believe (though perhaps falsely) that the action in question counts as contributing to the goal invoked in (1). Now, where p is any proposition, if it is reasonable for a fictional character in a given context to believe that p, then it would be reasonable for anyone in the same context, relevantly specified, to believe that p, whether that person is *real or fictional*. Another way of putting this is that whether the person concerned is real or fictional is not, on its own, relevant to what counts as reasonable for such a person to believe in a given context. What *is* relevant is the rest of the person's mental states at the time (or, in the fictional case, what other mental states the person is fictionally depicted as having): her beliefs, desires, intentions, hopes, fears, and so on. This reflects a point familiar from the work of Donald Davidson: one's judgment that it is reasonable for a given agent to have a belief depends on one's also being able to attribute to her a whole network of other intentional states consistent with that belief (1980b: 221). To be capable of thinking of another as having propositional attitudes at all presupposes that her intentional mental states largely

cohere (Davidson 1985: 245). If this is right, and if we assume that thinking of fictional characters involves thinking of them as having propositional attitudes (as it does when thinking of real people), then our ability to think of them as having propositional attitudes depends on our interpretations of them being governed by the same standards which govern our interpretations of rational agents in general. These standards apply to real and hypothetical or fictional agents alike.

So successful execution of (2) tells us, in effect, that some actual (rational) human being could believe the action in question counts as contributing to the goal invoked in (1). Thus, the successful execution of (1) and (2) together, insofar as it amounts to showing the substantive intelligibility of a given action in terms of certain motivation(s), provides a demonstration that it is possible for someone (a real person, not just a fictional character) to be motivated to act in just this way. This, I suggest, brings us genuine psychological understanding of a potentially new and important kind.

An instructive contrast may be made with what I earlier called 'weak intelligibility.' Recall the example of a character desiring a saucer of mud, with which to rub mud in her face in order to make her face the same color as the wardrobe. Assuming this is weakly but not substantively intelligible as a motivation (insofar as it does not readily fall under any desirability characterization), simply seeing that these mental states cohere, in that one action obviously counts as a means to the next, does not thereby bring with it any knowledge about the possibility of their combination within some real human being. Until we can see why someone might want to make her face the same color as a wardrobe, and not just 'why' in the sense of some further, equally incomprehensible reason being posited (e.g. that 'she wanted to look like a tree'), but 'why' in the sense of it being shown why *this* might count as a desirable thing to do, we are none the wiser as to whether some real person might actually be motivated in this way. In contrast, a judgment that a given fictional action is substantively intelligible by virtue of its connection to some desire or moral judgment that is intelligibly linked to some further desirable end, brings with it the knowledge that this is a possible way for real human beings to be motivated. To read about *Beloved*'s Sethe or *Lolita*'s Humbert and to understand each of them is to know that some real human being could be motivated to act as each, fictionally, acts.

This knowledge is not acquired by induction – that is, we do not need to find at least one other instance of someone who is motivated as Humbert is 'in the great world' before we can assent to the possibility of this combination of desire and action. The knowledge is acquired just by our finding the action in question substantively intelligible. The substantive intelligibility of a given action does not at all count as *partial* evidence that some agent could act in that way; rather, it establishes this conclusively. Hence it is not true that "art never confirms its truths" (Stolnitz 1992: 196). We have found confirmation of the observation, made against Carroll, that "our beliefs about moral psychology can be refined and changed in abstraction from knowledge of brute empirical facts" (Conolly and Haydar 2001: 121). Fiction is then potentially a 'self-sufficient' source of psychological understanding; and so there is no reason to

deny either that such understanding can be substantially new to the reader or that it can be genuinely interesting and complex.

Earlier I admitted that it is possible for a reader to be brought to see a given goal as desirable ceteris paribus whilst continuing to find unintelligible a character's prioritization of that goal over other goals, and her indifference to the (from the reader's perspective) undesirability of the goal's consequences. Does this threaten my claim that fiction can provide conclusive knowledge of psychological possibilities? I don't think so. Even in those cases where a goal is shown to be desirable yet unintelligible, insofar as it is prioritized by a character over other apparently more valuable goals, something can be learnt, namely, that the goal can be intelligibly pursued at all. This is so even though it may still be unclear how the goal can be intelligibly pursued in the precise context in which the character pursues it.

It is a consequence of my view that certain works of fiction are on a par with diaries or autobiographies with respect to what they can reveal to the reader about possible motives for action. Obviously, diaries and fictional works differ in that only the former can provide information about the motives upon which agents *have* acted. Here I suppose one might insist that information about possible motives for acting is uninteresting, unless it is *also* information about some historical agent's actual motives, as is found in diaries. Stolnitz intimates as much when he criticizes fiction for not being able to show that a given reason, explored in a fiction, could be 'the primary' or 'inevitable' cause of a real action (1992: 195). However, it is not clear why we should accept this bias towards the actual. Whether a given rationalization is of a fictional action or one that has actually occurred seems irrelevant to whether we find it illuminating or not. Accounts of the purported motives of real agents can be profoundly unilluminating: consider the *prima facie* unintelligibility of a wife-beater's account of his activities, who claims that he acts thus 'because he loves her.' Meanwhile, although the discovery that *The Diary of Anne Frank* is a fiction (as is sometimes suggested) might lessen the power of the book in certain ways, it seems obvious that the force of the lessons one has learned, from that work about what was *possible* for a adolescent growing up in Holland under the Nazis – how she might think, feel, desire, and act in such a situation – would not be undermined.

Alternatively, one might accept that fiction can provide us with psychological insights in the form of knowledge of possible motives for action, yet question the value of these insights on the following grounds. One could agree it is *possible* that someone might be motivated to kill a child in the light of such motives as Sethe has; nevertheless, given the specificity of Sethe's epistemic situation (including references to the background slavery, her specific perception of what she has suffered in the past, her beliefs about the supernatural, and so on), it is highly unlikely that any person would be, *in actuality*, so motivated. This is reminiscent of one horn of a dilemma intimated by Stolnitz: either the psychological portraits offered by fictions are so detailed in their specifications of fictional characters and events that

they cannot be applied to the 'real world;' or they are shorn of such elements, in which case we are left only with truisms (1992: 193-4).

We can reject the first horn of this dilemma, however. I have insisted that often part of what it is to explain an otherwise unintelligible fictional action is to redescribe it so as to give it a place in relation to a pattern of attitudes and beliefs which thereby reveal it as related to the pursuit of a goal which the reader judges, all things being equal, as desirable. For this to be possible, the action and the attitudes which motivate that action, no matter how specifically described, must also be describable in more general terms. Whether a given fictional action is judged to have a recognizable point will depend, not on features of the situation described at the most specific or particular level, but on whether that action possesses more general, abstract features repeatable elsewhere. For instance, a reader of *Beloved* can understand Sethe's murder of her child only if that reader sees the murder as motivated by a desire to save, not just those children *full stop*, but those children because they are to Sethe what is "precious and fine and beautiful" (Morrison 1987: 163). These are characteristics which might (though not, perhaps, for Sethe) be true of other people in other circumstances.

Additionally, as we have seen, an action is made intelligible by showing how the agent's belief that the action will fulfill some goal of hers is itself intelligible; that is, by showing what features of the action might be seen by the agent as contributing to, or consisting in achievement of, the goal in question. For this, too, the action must be describable in terms of relatively general features relevant not only to this situation in particular, but elsewhere.

No psychological information conveyed by a work of fiction is so indelibly tied to the context of the fiction that it cannot be made potentially relevant to real human beings. Therefore, worries of the sort voiced by Stolnitz are unfounded.

Finally, one might object that any knowledge acquired in this way from fictions is not *propositional* knowledge, strictly speaking, and so does not falsify the claims of those who argue that fiction cannot give us psychological insight *of a propositional sort*. I am not sure what this objection might rest on, other than a prejudice that the proper objects of propositional knowledge are confined to only those propositions which can be easily or briefly conveyed or assimilated. Here, I do not think we should overstate the requirements on what counts as propositional knowledge. A piece of propositional knowledge can take the form of a collection of many propositions, including propositions of extremely complex form and/or content. This being so, there seems to be no impediment to analyzing, as propositional in form, psychological information about the connections between a given act and the possible motivations behind it – information that one can clearly acquire from reading fiction. Granted, the kind of psychological understanding of the possibilities of human mental life one gains through reading fiction is not aphoristic in form, and can take much time and effort to extract from a given fiction, but this does not entail that it is not propositional. It means only that its characterization can be very complicated. That is, I suggest, as we would wish.

Acknowledgments: I wish to thank Peter Goldie for his very helpful comments on a draft of this essay. I also thank two anonymous referees.